Exceptional service in the national interest

Sandia National Laboratories

Uncertainty Treatment in Performance Assessment for a salt repository - the Sandia Approach

Cédric J. Sallaberry

3rd US/German Workshop on Salt Repository Research, Design and Operations

Albuquerque, NM, USA

October 9-10, 2012

SAND2012-8036P

Sandia National Laboratories

DBE TEC DBE TECHNOLOGY GmbH

PTKA Project Management Agency Karlsruhe within Karlsruhe Institute of Technology

U.S. DEPARTMENT OF ENERGY

NNSA National Nuclear Security Administration

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND NO. 2012-XXXXP

What's a performance Assessment

Every PA starts with the same four questions

Q1: What can happen?

Q2: How likely is it to happen?

Q3: What are the consequences if it does happen?

Q4: How much confidence do you have in the answers to the first three questions?

Basic concept (1/4)

From the question to the mathematical characterization

EN1: Probabilistic characterization of what can happen in the future

- Answers first two questions
- Provides formal characterization of aleatory uncertainty

EN2: Mathematical models for predicting consequences

- Answers third question

EN3: Probabilistic characterization of uncertainty in PA inputs

- Basis for answering fourth question
- Provides formal characterization of epistemic uncertainty

Basic concept (2/4)

from mathematical characterization to implementation

Sampling-Based approach

EN3

1. Characterization of uncertainty in **e** (i.e., definition of $D_1, D_2, \ldots, D_{nE}$)

2. Generation of sample from **e** (i.e., generation of $\mathbf{e}_k$, $k = 1, 2, \ldots, nS$, in consistency with $D_1, D_2, \ldots, D_{nE}$)

EN2

3. Propagation of sample through analysis (i.e., generation of mapping $[\mathbf{e}_k, \mathbf{y}(\mathbf{a}|\mathbf{e}_k)]$, $k = 1, 2, \ldots, nS$)

Defined in
EN1

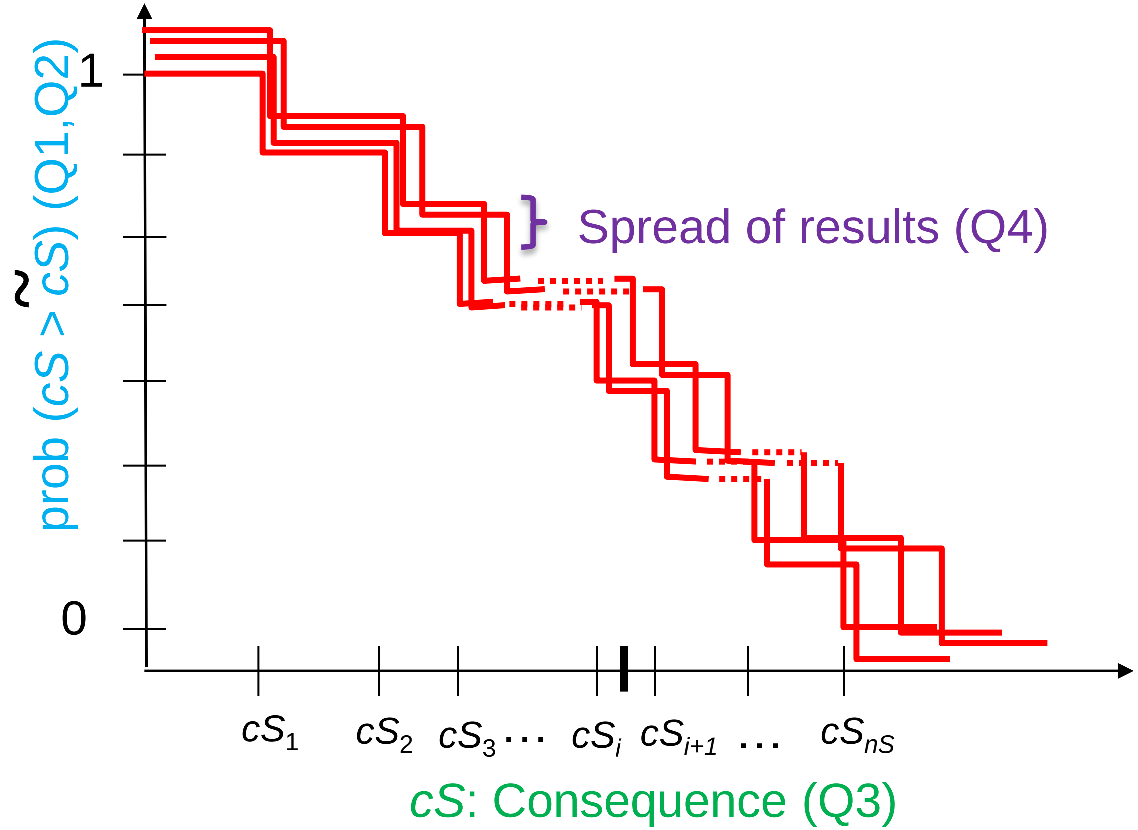
4. Estimate of uncertain results and statistical (aka uncertainty) analysis (i.e., approximations to the distributions of the elements of **y** obtained from $\mathbf{y}(\mathbf{a}|\mathbf{e}_k)$, $k = 1, 2, \ldots, nS$)

5. Determination of sensitivity analysis results (i.e., exploration of the mapping $[\mathbf{e}_k, \mathbf{y}(\mathbf{a}|\mathbf{e}_k)]$, $k = 1, 2, \ldots, nS$)

Basic concept (3/4)

from implementation to results (uncertainty analysis)

- Risk expressed as a <u>family</u> of complementary cumulative distribution functions (CCDFs)



Basic concept (4/4)

from implementation to results (sensitivity analysis)

- Involves exploration of mapping [$\mathbf{e}_k$, $\mathbf{y}(\mathbf{a}|\mathbf{e}_k)$], $k = 1, 2, \ldots, nS$
- Available techniques for sensitivity analysis
  - Examination of scatterplots and cobweb plots
  - Correlation and partial correlation analysis
  - Stepwise regression analysis
  - Rank transforms to linearize monotonic relationships
  - Nonparametric regression: Loess, additive models, projection pursuit, recursive partitioning
  - ….

WIPP as an Example (1/4)

mathematical model

EN1

- Defined by probability space $(\mathcal{A}, \mathbb{A}, p_A)$,

  $\mathcal{A}$ = {**a**: **a** is a possible 10,000 yr sequence of occurrences at WIPP}
- Extensive review of possible disruptions at the WIPP led to drilling intrusions and potash mining being the only occurrences incorporated into the definition of $\mathcal{A}$

EN2

- Function **y** = **f**(**a**)
- Models represented by **f** include: Systems of algebraic equations, Systems of ordinary differential equations (ODEs), Systems of partial differential equations (PDEs), Algorithmic procedures
- Processes modeled include: Material deformation, Corrosion , Microbial gas generation, Two-phase fluid flow, Pressure-induced fracturing, Regional groundwater flow, Radionuclide transport in flowing groundwater

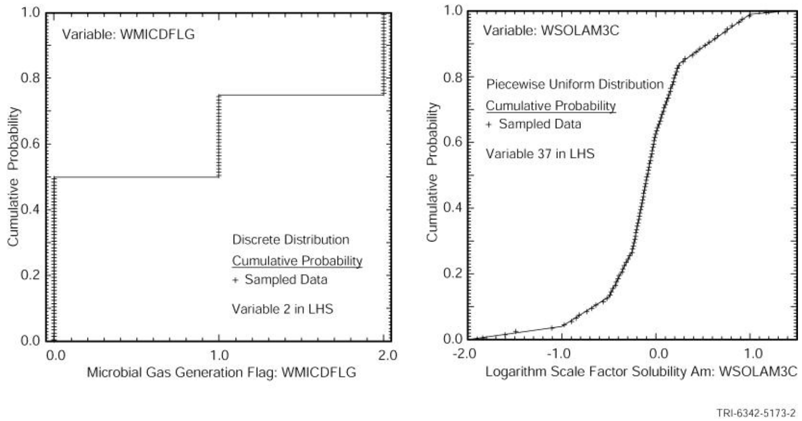
EN3

- Basic ideas underlying definition of EN3: Acceptable definitions for EN1 and EN2 result if appropriate values are assigned to quantities represented by $\mathbf{e}=[e_1,e_2,\ldots e_{NE}]$. Many possible values of varying levels of plausibility exist for **e.** Uncertainty in appropriate value to use for each e_j can be characterized by a distribution D_j
- EN3 defined by probability space $(\mathcal{E}, \mathbb{E}, p_E)$ derived from $D_1, D_2, \ldots, D_{nE}$

WIPP as an example (2/4)

implementation

- 57 uncertain quantities considered in WIPP 1996 PA, each associate with a distribution, including *WGRMICI* (Gas generation rate due to microbial degradation of cellulose under inundated conditions), *BHPRM* (Borehole permeability), *WTAUFAIL* (Shear strength of waste)…

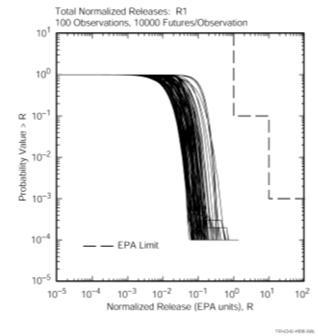


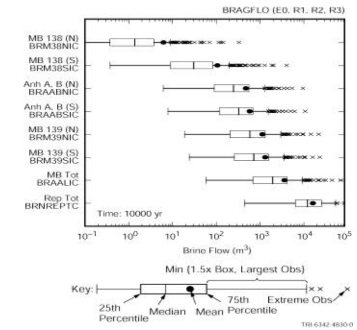
- Sample generated as 3 independent replicated samples of size 100 each
- Iman/Conover technique used to control correlations
- Of course, development of **f** , which is the most complex part.

WIPP as an example (3/4)

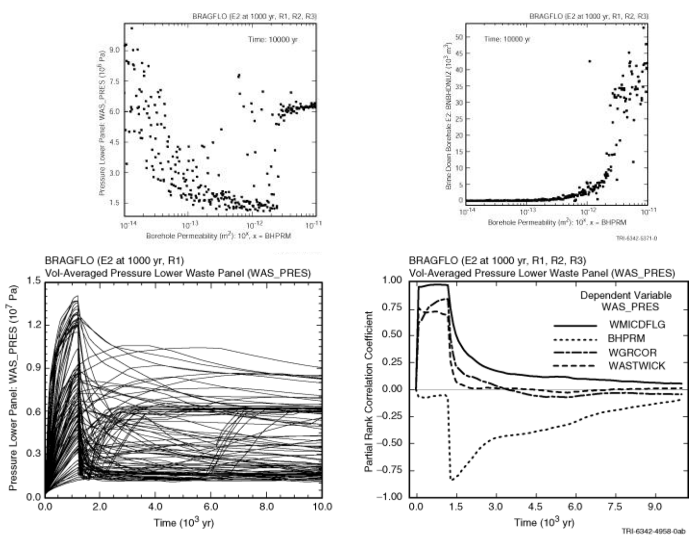
uncertainty and sensitivity analyses

- WIPP Results as distributions of CCDFs or time dependent expected values
- Conditional results presented as Box plots





- Scatterplots, partial correlation coefficients, and stepwise regressions were used to analyze the influence of inputs uncertainty over selected outputs uncertainty.



Pressure at 10^4 yr

Step[a]	Variable[b]	SRC[c]	R^{2}[d]
1	*WMICDFLG*	0.718	0.508
2	*HALPOR*	0.466	0.732
3	*WGRCOR*	0.246	0.792
4	*ANHPRM*	0.129	0.809
5	*SHRGSSAT*	0.070	0.814
6	*SALPRES*	0.063	0.818

[a] Steps in stepwise analysis.
[b] Variables listed in the order of selection in regression analysis.
[c] Standardized regression coefficients (SRCs) for variables in final regression model.
[d] Cumulative R^2 value with entry of each variable into regression model.

WIPP as an example (4/4)

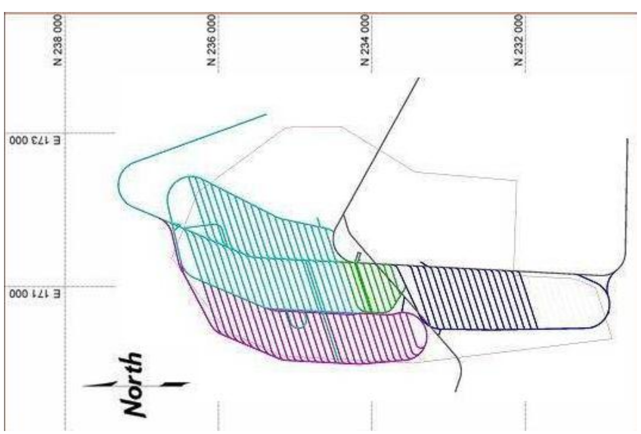
conclusion

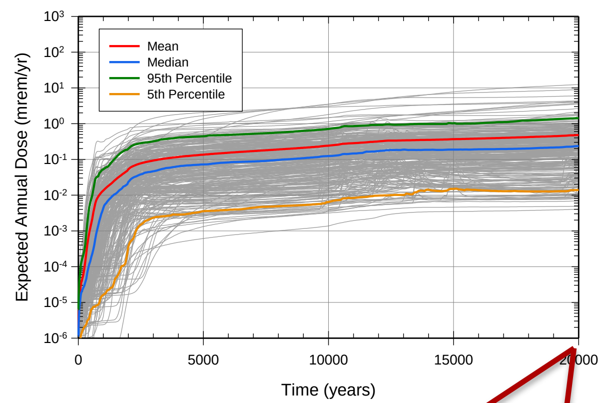
- Certification completed in 1996 (16 years ago)
- Most of the approaches used (sampling based method, rank regression …) are still valid today
- Remains a reference in term of Successful Probabilistic Assessment for underground repository
- But in order to remain an reference in radioactive waste disposal, Sandia (via center 6200) works into maintaining the past knowledge and improving or working on new concepts and techniques.
- Following slides will present Improved and new techniques successfully used at Sandia, complementing the original expertise

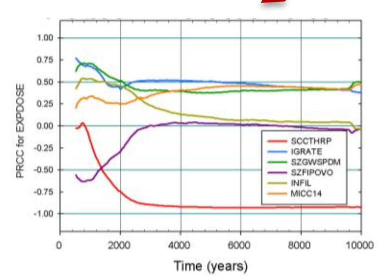
Different repository: same approach

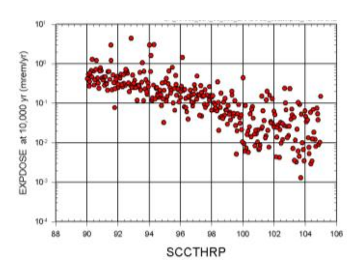
the Yucca Mountain example

- Different repository concept (fractured media, high level waste, engineered barrier system with drip shield)
- Different regulatory requirements
- But same approach has been used with success to estimate uncertainty over expected doses and perform sensitivity analysis







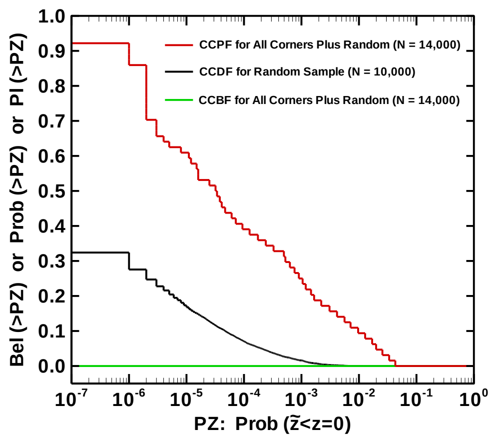


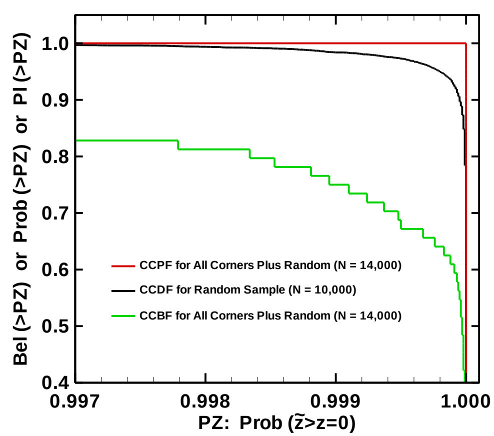
10 years after, we can still do it … and in color !

New approaches and methods

Alternative representations of (epistemic) uncertainty

- Sometimes, uncertainty is not well-enough defined to use probability without adding some subjectivity.
- Alternative representations of uncertainty (such as Evidence theory, Possibility theory, Fuzzy sets, Interval analysis) allow to relax some of the constraints inherent to probability theory and avoid such addition
- With current computational capabilities, Sandia showed that it is possible now to use in real-world problem such new representations
- CCDF over epistemic uncertainty are then defined with a range delimited by Plausibility (in red in graph below) and Belief (in green below)



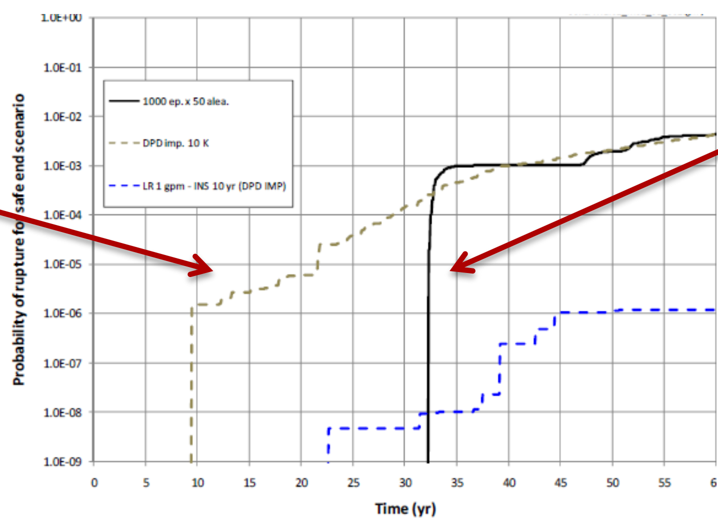


New approaches and methods

New sampling and optimizations techniques

- In term of sampling, LHS is still considered as the sampling of choice for many problem as it informs where most of the uncertainty has an effect
- The original code developed by Ron Iman in 1980 has been continually improved since
- **Sometimes, LHS may not be the method of choice**: when estimated probabilities are pretty low, when a particular area of interest needs to be oversampled …
- Other techniques such as **importance sampling and optimizations** are then more appropriate and efficient
- Sandia develops and maintain a freeware optimization toolkit that can be plug to external code and run optimization and/or sampling on the code. Its codename is DAKOTA

Importance sampling with same sample size is a lot more accurate for low probabilities events



Classical LHS missed low probability unless sample size is increased

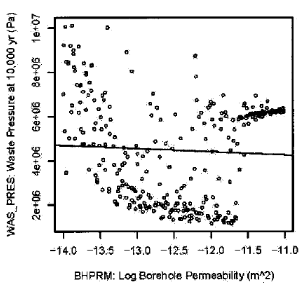
Results from xLPR project for US-NRC

New approaches and methods

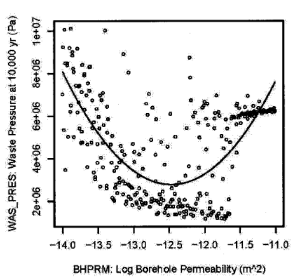
New regression techniques (1/2)

- Rank regression was a powerful technique that captured most of the uncertainty for WIPP and Yucca Mountain sensitivity Analysis
- However, it did not capture non monotonic influence (fig. A) and conjoint influence of two parameters
- New regression techniques, coupled with Sobol variance decomposition has been successfully used in order to capture more complex influence that sometimes happen in complex models and are hard to recognize. Techniques such as quadratic regression (fig. B), recursive partitioning (fig. C), MARS and ACOSSO (fig. D) have been used to capture more complex relationship between input and output.

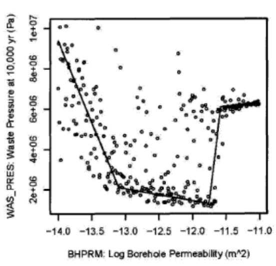
A



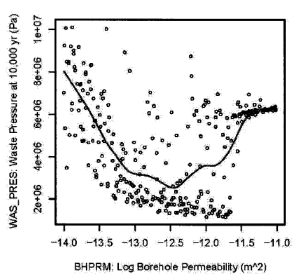
B



C



D



New approaches and methods

New regression techniques (2/2)

- These techniques have been also successfully used to capture conjoint influence that was missed by regression into additive models

	Rank Regression			Quadratic			Recursive Partitioning			MARS		
Final R^2	0.24			0.61			0.60			0.60		
Input	**R^2 inc.**	**R^2 cont.**	**SRRC**	**S_i**	**T_i**	**p-val**	**S_i**	**T_i**	**p-val**	**S_i**	**T_i**	**p-val**
CYSIGA.1..	0.04	0.04	-0.11	0.00	0.03	0.56	0.00	0.00	1.00	0.05	0.17	0.14
EFFTHR.1..	0.08	0.04	-0.11	0.06	0.13	0.26	0.06	0.09	0.23	0.00	0.48	0.00
GSHFAC.2..	0.11	0.03	0.11	0.01	0.24	0.02	0.00	0.13	0.04	0.02	0.11	0.25
SRVLAMB	0.13	0.03	-0.09	0.00	0.30	0.00	0.12	0.36	0.00	0.01	0.26	0.02
CWASH1.	0.16	0.03	0.08	0.00	0.86	0.00	0.26	0.91	0.00	0.13	0.34	0.14
SRVOAFRAC	0.18	0.02	-0.07	0.01	0.00	1.00	---	---	---	0.06	0.16	0.62
RRDOOR	0.19	0.01	0.10	0.00	0.00	1.00	---	---	---	0.01	0.32	0.00
PROTIN.2..	0.20	0.01	0.05	---	---	---	0.00	0.00	1.00	0.02	0.43	0.00
EFFTHR.2..	0.21	0.01	-0.05	---	---	---	0.00	0.30	0.00	0.02	0.00	1.00
CSFACT.3..	0.22	0.01	0.04	---	---	---	---	---	---	0.00	0.10	0.26
EIFACB.3..	0.22	0.01	-0.04	---	---	---	---	---	---	---	---	---
SC1131_2	0.23	0.01	0.03	---	---	---	0.03	0.15	0.05	0.00	0.20	0.03
DLTEVA.12..	0.23	0.01	-0.04	0.04	0.46	0.00	---	---	---	---	---	---
SRVFAILT	0.23	0.00	0.03	0.00	0.00	1.00	0.00	0.00	1.00	---	---	---
EITHRE.3..	0.24	0.00	-0.03	0.00	0.00	1.00	---	---	---	0.00	0.20	0.03
DLTEVA_5.2..	---	---	---	0.00	0.24	0.01	---	---	---	---	---	---
CSFACT.2..	---	---	---	0.00	0.27	0.09	0.00	0.48	0.00	0.02	0.02	0.56
PROTIN.3..	---	---	---	0.00	0.07	0.32	0.00	0.00	1.00	---	---	---
PROTIN.1..	---	---	---	0.01	0.00	1.00	---	---	---	0.01	0.00	1.00
VDEPOS.1..	---	---	---	0.00	0.00	1.00	0.04	0.09	0.07	---	---	---
DLTEVA_5.11..	---	---	---	---	---	---	0.02	0.00	1.00	---	---	---
EFFACB.1..	---	---	---	---	---	---	0.01	0.00	1.00	---	---	---
DLTEVA.13..	---	---	---	---	---	---	0.00	0.00	1.00	---	---	---
GSHFAC.3..	---	---	---	---	---	---	---	---	---	0.00	0.00	1.00
	---	---	---	---	---	---	---	---	---	---	---	---
	---	---	---	---	---	---	---	---	---	---	---	---
	---	---	---	---	---	---	---	---	---	---	---	---

In this example, the conjoint influence of variable CWASH1 was completely missed by Rank Regression

Examples results from SOARCA UA project for US-NRC

Conclusion

- Sandia has shown capability to complete a complex PA for certification of radioactive waste repository in Salt in the past

- The expertise in the area has been maintained as many of the techniques used are still valid and powerful

- Since then, Sandia has built upon this basis to expand its expertise and bring new and improved techniques to enrich its capabilities

- These techniques are not only theoretical but **used** in **real** projects with success in other activities.

- Our desire is now to regroup all these techniques and apply them again on Repository science to demonstrate their advantages